

Whole-genome genotyping of grape using a panel of microsatellite multiplex PCRs

Belkacem Zarouri • Alba María Vargas • Laura Gaforio •
María Aller • María Teresa de Andrés •
José Antonio Cabezas

Abstract The use of microsatellite markers in large-scale genetic studies is limited by its low throughput and high cost and labor requirements. Here, we provide a panel of 45 multiplex PCRs for fast and cost-efficient genome-wide fluorescence-based microsatellite analysis in grapevine. The developed multiplex PCRs panel (with up to 15-plex) enables the scoring of 270 loci covering all the grapevine genome (9 to 20 loci/chromosome) using only 45 PCRs and sequencer runs. The 45 multiplex PCRs were validated using a diverse grapevine collection of 207 accessions, selected to represent most of the cultivated *Vitis vinifera* genetic diversity. Particular attention was paid to quality control throughout the whole process (assay replication, null allele detection, ease of scoring). Genetic diversity summary statistics and features of electrophoretic profiles for each studied marker are provided, as are the genotypes of 25 common cultivars that could be used as references in other studies.

Keywords SSR • Null allele • *Vitis vinifera* • Multiplex PCR • Genetic diversity

Introduction

During the last decades, microsatellites, also known as simple sequence repeats (SSRs), have been widely used for different genetic studies in plants (Kalia et al. 2010). Their popularity is related to their monolocus and multiallelic features, codominant inheritance, and high reproducibility (Rafalski and Tingey 1993; Powell et al. 1996; Schlötterer 2004). In addition, they are hypervariable, the data are easily comparable between laboratories, and they require basic equipment accessible for small- to medium-sized laboratories. In spite of these advantages, microsatellite markers present some limitations: mainly the high cost and time-consumed per individual SSR assay (Guichoux et al. 2011). In order to increase cost and labor efficiency of SSR analysis, multiplex PCR approaches combined with fluorescence-based analysis and semi-automated allele calling have been developed (Butler 2005b; Missiaggia and Grattapaglia 2006; Blacket et al. 2012). These strategies allow to simultaneously amplify more than one locus in a single reaction using multiple primer pairs. Multiplex PCR protocols have been developed and successfully applied for different uses in many species, such as the detection of colorectal tumors in humans (Patil et al. 2012), monitoring of rat strains (Bryda and Riley 2008), genetic characterization of different plant species (Jewell et al. 2010; Postolache et al. 2013; Drašnarová et al. 2014), identification of selfed progenies in switchgrass (Liu and Wu 2012), or to facilitate systematic and rapid genetic mapping in soybean (Sayama et al. 2011). Multiplex PCR design requires a priori knowledge of the range of allelic sizes for each marker in order to avoid amplicons overlapping and the optimization of PCR conditions to obtain a balanced amplification of all the targeted fragments (Butler 2005a). Another limitation for the use of microsatellites is the eventual presence of null alleles that challenges particularly population and linkage disequilibrium-based studies (Callen et al. 1993). A

microsatellite null allele is any allele that cannot be amplified via PCR, usually due to mutations in the primer binding sites (Dakin and Avise 2004), and may lead to misinterpreting the results by: (1) considering a PCR failure when the null allele is present in homozygous state; (2) classifying a heterozygote individual as homozygote when the null allele is in heterozygous state (Wagner et al. 2006).

Grapevine is one of the most important perennial fruit crops in the world (<http://faostat.fao.org>). The cultivated monoecious form (*Vitis vinifera* L. ssp. *sativa*) was domesticated from the wild dioecious still existing form (*V. vinifera* L. subsp. *sylvestris* (Gmelin) Hegi) in the Near East about seven to eight thousands of years ago (This et al. 2006; Myles et al. 2011). Following domestication, thousands of cultivars derived from spontaneous or controlled crosses, but also from somatic variation (Torregrosa et al. 2011), have been selected and spread by vegetative propagation throughout the world, from temperate to tropical climates (Bouquet 2011). The highly heterozygous diploid genome of grapevine has a polyploid origin (Jaillon et al. 2007). Its relatively small size (475 Mbp and $2n=38$ chromosomes; Lodhi and Reisch 1995) has facilitated a significant progress in grapevine genomics, being the publication of the genome sequence in 2007 the most important one (Velasco et al. 2007; Jaillon et al. 2007). The availability of the grapevine genome sequence combined with the advent of cheaper and high throughput single nucleotide polymorphism (SNP) genotyping strategies (Gupta et al. 2008; Davey et al. 2011) were expected to shift the tools of genetic studies in grapevine. However, microsatellites are still the predominant markers contributing to the current knowledge of genetic determinism of the major grapevine traits (Mejía et al. 2011; Huang et al. 2012; Duchêne et al. 2012; Karaagac et al. 2012; Doligez et al. 2013; Battilana et al. 2013; Grzeskowiak et al. 2013; Ban et al. 2014; Correa et al. 2014).

Microsatellite repeats are abundant and diverse in the grapevine genome (Thomas et al. 1993). This has allowed the identification of hundreds of them and the design of primers for their analysis throughout the last two decades (Thomas and Scott 1993; Bowers et al. 1996, 1999; Sefc et al. 1999; Scott et al. 2000; Lefort et al. 2002; Decroocq et al. 2003; Arroyo-García and Martínez-Zapater 2004; Di Gaspero et al. 2005; Merdinoglu et al. 2005; Cipriani et al. 2008; Huang et al. 2011); up to 1,079 *V. vinifera* SSR probes can be currently found at the NCBI database (<http://www.ncbi.nlm.nih.gov/probe>). They have been used for many purposes, including the following: germplasm characterization and pedigree reconstruction (Sefc et al. 2009); construction of linkage maps (Cipriani et al. 2011); identification and mapping of quantitative trait loci (QTLs) (Welter et al. 2011) and marker assisted selection (Töpfer et al. 2011). However, to our knowledge, the unique study focused on the extensive design of multiplex PCRs for grapevine genotyping was reported by Merdinoglu et al. (2005). In

that study, 125 SSRs selected regardless of their position in the genome were grouped in 46 multiplex PCRs with up to three loci per multiplex. For fragment analysis, another round of multiplexing was needed to develop 22 multi-loading sets combining up to four multiplex PCRs per load. Additionally, there are a few published studies in which multiplex PCR protocols have been developed and used for cultivar identification and/or germplasm characterization (Ibáñez et al. 2009; Laucou et al. 2011; Moreno-Sanz et al. 2011; Migliaro et al. 2012). Any case, none of these tools allows to carry out genome-wide studies in grapevine, such as genetic mapping or SSR-assisted backcrossing as proposed by Herzog et al. (2013).

The use of microsatellite markers in genetic mapping or genome scanning-based breeding involves the following: the selection of markers based on their chromosomal position; the determination of their informativeness in the targeted mapping or breeding populations; and the genotyping of each individual with those markers. This could be a very expensive and time-consuming process, but it could be optimized using a carefully designed multiplex PCR approach. Our objective in the present study was to develop a panel of multiplex PCRs allowing to genotype microsatellite markers covering most of the *V. vinifera* genome. With this aim in mind, we first selected a set of microsatellite loci evenly distributed along grapevine chromosomes, searching also for new SSRs in uncovered regions; then, designed and optimized the multiplex PCRs; and finally, validated them using a set of grapevine accessions which represent a large extent of the existing *V. vinifera* genetic diversity, including several pedigrees to test allelic inheritance. Furthermore, this genotyping has provided a thorough knowledge about allelic diversity and associated parameters for each locus, which may be used to select the most interesting markers for future studies.

Materials and methods

Plant material

Two sets of plant material were used in this study (Online Resource 1: Table S1): (1) the “testing collection,” consisting of seven grapevine cultivars (Airén, Cabernet Sauvignon, Cardinal, Chardonnay, Crimson Seedless, Flame Seedless, and Italia); (2) the “validation collection,” consisting in a large set of 207 non-redundant grapevine genotypes with different uses and origins, selected to represent the genetic diversity contained in the 1,852 *V. vinifera* accessions that are maintained at the germplasm bank of El Encín (Alcalá de Henares, Madrid, Spain; <http://www.madrid.org/coleccionvidencin/>). The selection of the accessions was carried out using the genotypic information of the 26 SSRs included in multiplex PCRs Mx01 and Mx02 (Table 1),

Table 1 (continued)

Multiplexes	Marker 1	Marker 2	Marker 3	Marker 4	Marker 5	Marker 6	Marker 7	Marker 8	Marker 9	Marker 10	Marker 11	Marker 12	Marker 13	Marker 14	Marker 15
Mx40	VVIQ25-2	VVIB22-2	VVIV37	VMC4H2	VVIS70-2				RTAChr9-11b						
Mx41	VVIM25	RTAChr3-17a	RTAChr3-19b	RTAChr4-9a	RTAChr13-20a	RTAChr15-4a	VVIV67	VVIP08							
Mx42	RTAChr1-14a	RTAChr3-15a	FAM65	VVIQ52	FAM62-2	RTAChr15-3b	RTAChr16-6b	RTAChr17-15a							
Mx43	RTAChr2-11a	UDV-021-2	VRZAG67	RTAChr13-13b	RTAChr15-1b	RTAChr15-7a	VMCIE11-2	RTAChr9-16b							
Mx44	VVIV36	RTAChr9-13a	VRZAG67	RTAChr13-16a	RTAChr17-15b	UDV-134-2									
Mx45	VVMD27	RTAChr7-9	FAM104	RTAChr11-11b	RTAChr13-23b										

referred to as “control loci” from now on, and the maximization strategy implemented in the software MStrat v4.1 (Gouesnard 2001). Furthermore, the validation collection includes 14 known trio pedigrees (two progenitors and one progeny, Online Resource 1: Table S1) involving 31 among the 207 accessions.tgroup

DNA extraction

Frozen young leaves of each accession were ground to a fine powder using the Mixer Mill MM300 grinder (Retsch, Haan, Germany) and liquid nitrogen. Total DNA was purified using the BioSprint 96 workstation and the BioSprint 96 DNA Plant Kit (Qiagen, Hilden, Germany) according to the manufacturer instructions. Extracted DNA was quantified using the ND-1000 spectrophotometer (NanoDrop, Delaware, USA).

Markers selection

Initially, 249 SSR markers were selected from the literature and genomic databases (NCBI: <http://www.ncbi.nlm.nih.gov/probe> and GENOSCOPE: <http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>; Online Resource 1: Table S2) based on their physical position (inter-marker distance approx. <3 Mbp according to the 12× version of the *V. vinifera* genomic sequence (GENOSCOPE), and the available information about their degree of polymorphism). Alternative primers were designed for 37 of them (the new marker was named adding “-2” at the end of its original name) for which amplification problems or evidences of existence of null alleles were detected (Online Resource 1: Table S2). Their genomic sequences were retrieved from the 12× version of the *Vitis* genome sequence by Blast search using as probes the sequences of the original primers. Then, new primers were designed in the flanking regions of the microsatellite repeats using the software Primer3 v4.0 (Rozen and Skaletsky 2000) and tested for PCR amplification. Additionally, primers for 22 new microsatellite markers were designed in silico in order to cover chromosomal regions for which no markers were available (Online Resource 1: Table S2). Genomic sequences at intervals of approximately 2 Mbp within the chromosomal gaps were retrieved from the 12× grapevine genome sequence and investigated using the WebSat software (Martins et al. 2009) to identify microsatellites and primer pairs for their amplification. The specificity of the candidate primer pairs was checked by aligning them against the genome sequence using Primer-BLAST (Ye et al. 2012). Finally, the selected primer pairs were submitted to amplification test as described in the next section.

Multiplex PCRs design

Allelic information for many selected SSRs was only available for the accession in which the marker was identified or a few accessions more. For that reason, an expected fragment-size range was established for each selected locus by adding 30 bp to both sides of the allelic range determined according to the bibliographic and database information. Then, they were organized in groups of at most three markers with non-overlapping fragments to be labeled with the same fluorescent dye: FAM, NED, PET, or VIC (Applied Biosystems, Foster City, CA, USA). After that, quadruplets combining four different dye groups were constructed to constitute the multiplex sets.

Amplification tests

Labeled primer pairs for each locus were tested by genotyping the testing collection (see “Plant material”) using single-locus PCR in order to check their functionality and to confirm/correct allelic ranges. Single-locus PCRs were performed in a total volume of 21 μ l containing 5 ng of DNA template, 1 \times PCR buffer, 2 mM MgCl₂, 0.2 mM of each dNTP, 0.2 μ M of each primer, and 1 U of Taq DNA Polymerase (Biotools, Madrid, Spain). Amplifications were carried out using the following thermocycling conditions: 1 cycle at 95 °C for 5 min; followed by 35 cycles at 94 °C for 30 s, 52 °C for 1 min (suitable for most of the loci and modified for the rest), and 72 °C for 1 min; and a final step of 72 °C for 30 min. PCR products were separated by capillary electrophoresis in an ABI3130 sequencer (Applied Biosystems, Foster City, CA, USA). The loading mixture contained 1 μ l of PCR product diluted 50 to 160 times (depending on the intensity of the amplified fragments on 2 % agarose gels), 0.1 μ l of internal size standard (GeneScan-500LIZ; Applied Biosystems, Foster City, CA, USA), and 14 μ l of Hi-Di Formamide (Applied Biosystems, Foster City, CA, USA). Prior to be loaded into the sequencer, the samples were denatured at 95 °C for 5 min. Raw data obtained by capillary electrophoresis were transformed into allelic sizes using the GeneMapper v4.1 software (Applied Biosystems, Foster City, CA, USA).

Multiplex PCRs optimization

Based on the results of the amplification tests, the initially proposed multiplex sets were confirmed or corrected to avoid marker overlapping and then subjected to multiplex PCR assays with the same testing collection. Initially, all the Multiplex PCRs were performed in a final volume of 15 μ l containing the following: 1 \times Multiplex PCR Master Mix (Qiagen, Hilden, Germany), 5 ng of template DNA, and an equimolar amount of 0.2 μ M of each primer pair. Subsequently, each multiplex PCR was subjected to several

cycles of optimization, modifying primer concentrations according to the signal intensity (peak height) of the amplification products (decreasing for the stronger ones and increasing for the weaker ones), until an acceptable equilibrium between the combined primer pairs was reached.

Four touchdown thermocycling programs (Don et al. 1991), coded as Tp-A, Tp-B, Tp-C, and Tp-D (Online Resource 1: Table S3), were used for PCR amplification, varying the number of cycles and/or annealing temperature. All of them included a 90-min at 72 °C final extension step in order to address the “Plus-A” artifacts in PCR products (Smith et al. 1995). Each multiplex was initially tested using the programs Tp-A (standard program) or Tp-B (suitable for primers with lower annealing temperatures). When amplification deficiencies were detected for any locus, the other programs were tested to identify the best performing one. Finally, the designed and optimized multiplexes were validated genotyping the validation collection.

Data checking

In order to minimize possible mistakes, several control points were included throughout the study: (1) sampling errors and possible contaminations during DNA extraction were checked by confirming the identity of the sampled accessions using the 26 control loci (Mix01 and Mix02), which had been previously used to genotype the complete germplasm bank of El Encín (Alcalá de Henares, Madrid, Spain); (2) to detect possible errors during DNA manipulation or electrophoresis related problems, one control locus was included in each multiplex PCR (Online Resource 1: Table S2); (3) contaminations during the amplifications were controlled using negative controls in each PCR; (4) to verify assay reproducibility, 20 accessions were genotyped twice; (5) PCR and/or sequencer loading were repeated for any sample that failed to generate detectable amplicons; (6) both allele definition (binning process) and genotyping (peaks selection) in GeneMapper v4.1 (Applied Biosystems, Foster City, CA, USA) were manually reviewed by two persons prior to generate the final genotypic table; (7) finally, allelic inheritance was checked for each locus using the 14 trio pedigrees included in the validation collection. When incompatibilities were found, allele calling was reviewed for the whole validation collection to discard mistakes in peak selection or data typing. If the incompatibility persisted, the amplification of the problematic marker was repeated using single-locus PCRs with different DNA polymerases (Standard DNA polymerase or *pfu* DNA polymerase (Biotools, Madrid, Spain)) and/or different thermocycling parameters. When the compatibility could not be recovered, the locus in question was suspected to present null alleles or unspecific amplification and new primer pairs were designed for it.

Data analysis

The genotypic dataset obtained for the validation collection was used to estimate the following parameters for each analyzed locus employing Cervus 3.0 software (Kalinowski et al. 2007): number of detected alleles/locus (K), observed heterozygosity (H_{obs}), expected heterozygosity (H_{exp}), polymorphic information content (PIC), probability of identity of unrelated individuals (PI), and estimated frequency of null alleles (F). In order to evaluate the quality of the amplification profiles of the studied SSRs, we established a panel of seven descriptors related to the stuttering patterns, “+A” effect, low heterozygote peak ratio, presence of artifacts, separation between adjacent fragments, multilocus patterns, and a global evaluation of the easiness of scoring of the marker. These descriptors were evaluated by visual examination of the electrophoretic profiles as displayed in GeneMapper v4.1 (Applied Biosystems, Foster City, CA, USA).

Gene density distribution over the grapevine chromosomes was estimated using the information of the annotated genes from the 12× whole genome sequence available at the GENOSCOPE database (<http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>). Gene density along each chromosome was represented as number of genes in consecutive chromosomal bins of 500 kbp.

Results

Amplification tests

Six loci (FAM18, FAM68, FAM96, FAM104, VVIB54, and VVIV22) out of the 249 selected initially detected only one fragment in the testing collection, suggesting that they could be monomorphic. Even though, they were included in the designed multiplex sets, since more alleles could be detected in the 207 accessions of the validation collection. Three of them were confirmed as monomorphic (FAM18, FAM104, and VVIB54), whereas the rest (FAM68, FAM96, and VVIV22) detected only one allele more in the validation collection. Even though these markers were maintained in the multiplexes since they can be polymorphic in genetic backgrounds of wider diversity. Eleven markers amplified more than two fragments per accession; hence, they were classified as multilocus (Online Resource 1: Table S2). Six of them (UDV-020, UDV-038, VMC3G7, VMCNG1E4-2, VMCN G1D3, and VMC4B7-2) were included in the designed multiplex PCRs and analyzed in the validation collection, and an attempt was done to score each marker as two different loci. However, it was not possible to discriminate between the amplified fragments corresponding to each locus because they were located in many cases in the same narrow fragment size interval (Fig. 1). Even though, they were maintained in the

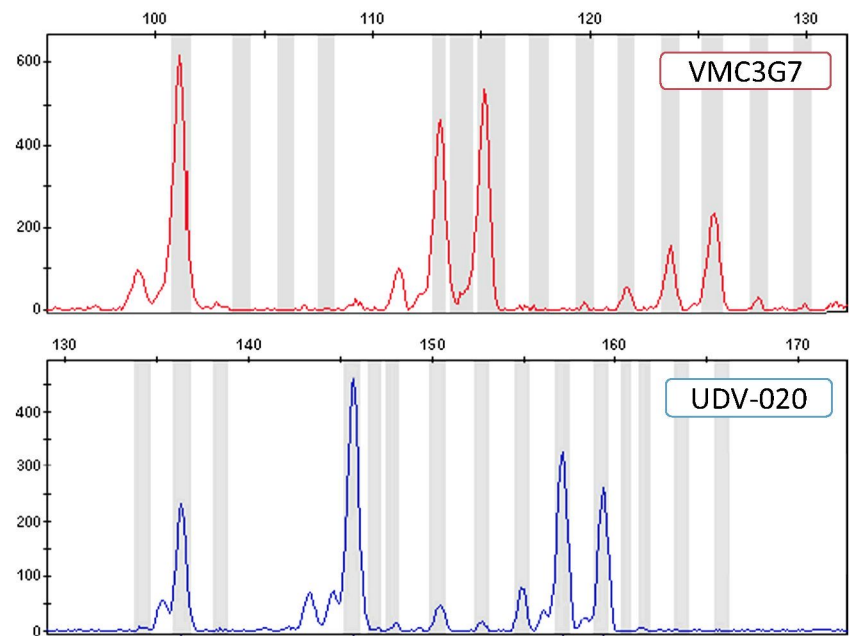
multiplexes since they can be useful for genetic mapping or breeding purposes in bi-parental populations, where allele screening could be less complicated. The locus VVS3 was also discarded because it generated irreproducible amplification profiles (repeated PCRs produced fragments of different sizes).

Multiplex PCRs design, optimization, and validation

Two hundred and thirty-six SSRs out of the 243 that passed the amplification tests were distributed in 34 multiplex sets (multiplexes from Mx01 to Mx34 in Table 1 and Online Resource 1: Table S2). The seven remaining SSRs could not be assigned to any multiplex because they lead to amplicons overlapping. PCR conditions were optimized by adjusting primer concentrations and thermocycling parameters. For each multiplex, two to eight (3.8 in average) optimization rounds were needed to reach a balanced amplification between the primer pairs combined in the same reaction. Final protocols involved the use of primer concentrations ranging from 0.04 to 0.60 μM (Online Resource 1: Table S2) and four touchdown-based thermocycling programs (Tp-A, Tp-B, Tp-C, and Tp-D; Online Resource 1: Table S3). Once optimized, these multiplex PCRs were used to genotype the validation collection. Inheritance analysis using the 14 trio pedigrees included in the validation collection revealed inconsistencies in 37 loci. Thirty-five of them showed incompatible genotypes in at least one of the pedigrees (Online Resource 1: Table S2) that were likely due to the presence of null alleles in the progenitors (i.e., observed “bb” offspring derived from “aa×bb” or “aa×bc” crosses). The unique exceptions were VMC5G7 and B003, which showed genotype inconsistencies in only one pedigree, but it did not fit the presence of null allele pattern. Those two cases may be the result of somaclonal mutations which are not rare in grapevine (Pelsy 2010). In an attempt to avoid null alleles and recover monolocus segregation, alternative primers were designed for 37 problematic markers (32 suspected of carrying null alleles, 4 multilocus (UDV-021, FAM62, VMC1E11, and UDV-134), and VVS3). Successful single-locus amplifications were obtained for 25 of them (Online Resource 1: Table S2).

SSR markers were unavailable in the literature for several chromosomal regions of the grapevine genome, leaving uncovered gaps of up to 10 Mbp according to the 12x.2 genome sequence (Fig. 2; Canaguier et al. 2014; <https://urgi.versailles.inra.fr/Species/Vitis/Data-Sequences/Genome-sequences>). Thus, we tried to develop new SSR markers in order to reduce those gaps. Although many microsatellite motifs were detected in the investigated genomic sequences, it was difficult to find specific primers for their amplification. As an example, in a segment of 1 Mbp from chromosomes 3 (between 13.5 and 14.5 Mbp) and chromosome 15 (between 2.5 and 3.5 Mbp), 91 and 153 microsatellites were identified, respectively. Among them, only 54 % yielded candidate

Fig. 1 Examples of electrophoretic profiles showing the possible amplification of multiple loci in one accession by the markers VMC3G7 and UDV-020. *Gray stripes* represent the bin set (possible alleles in the validation collection)



amplification primers, and their alignment against the grapevine genome sequence using Primer-BLAST revealed that most of them (90 and 81 %, respectively) had more than one matching site (i.e., susceptible to unspecific amplification). Finally, a total of 22 new SSR markers could be developed in those gaps and successfully amplified in single-locus PCRs (Fig. 2; Online Resource 1: Table S2).

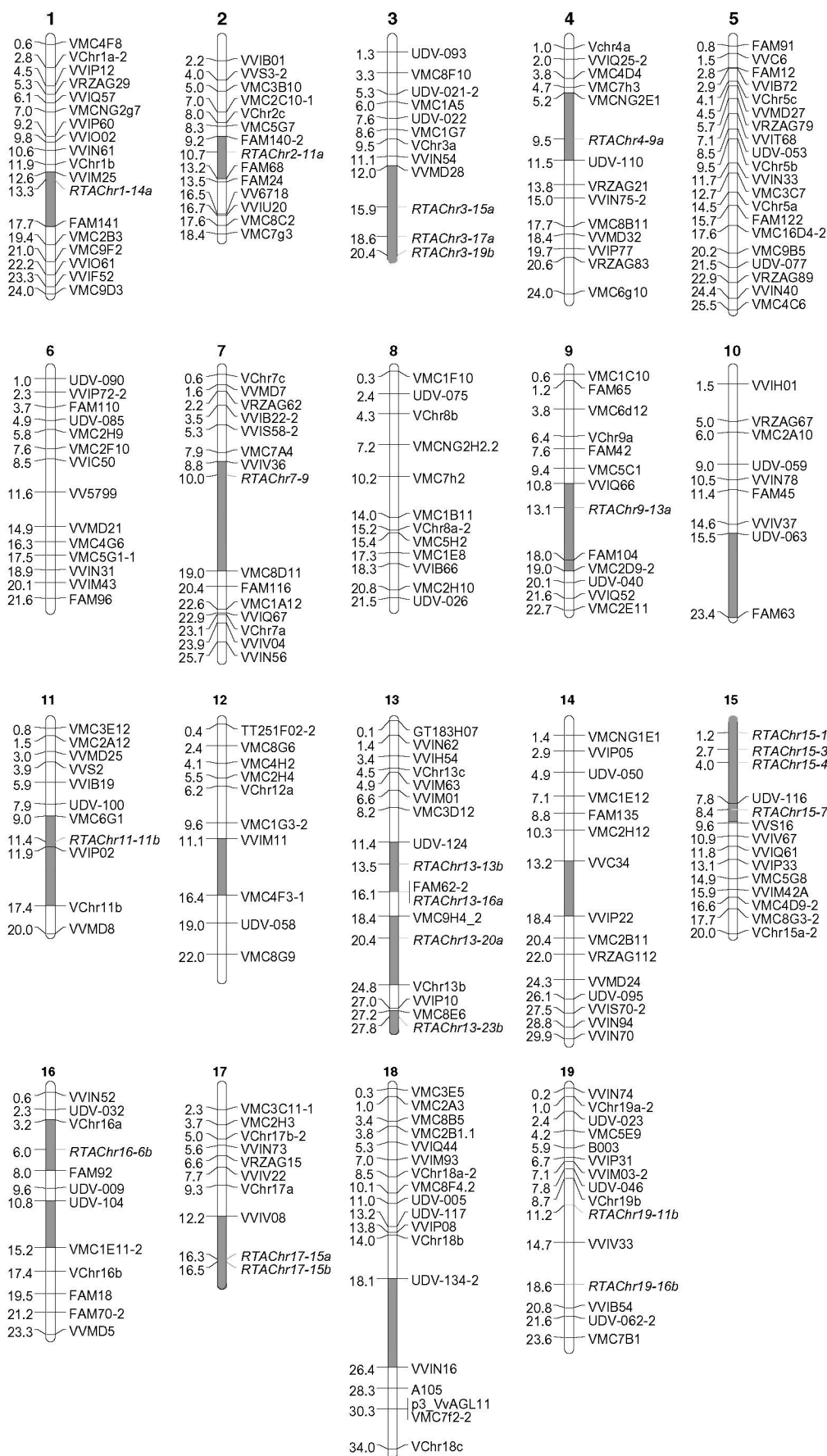
The 25 redesigned SSRs, the 22 newly developed, and the 7 remaining unassigned after the construction of the first set of multiplex PCRs were organized in a second set of 11 multiplexes (from Mx35 to Mx45 in Online Resource 1: Table S2). These multiplexes were then optimized and used to genotype the validation collection. The analysis of allelic inheritance in the trio pedigrees revealed that 3 out of the 25 redesigned markers and 10 out of the 22 new ones showed incompatible genotypes reflecting the presence of null alleles (Online Resource 1: Table S2).

In summary, 290 SSR markers (243 selected initially, 25 redesigned and 22 developed in this study) distributed homogenously along the 19 chromosomes of grapevine (Fig. 2) were successfully organized in 45 multiplex PCRs with an average of 7.31 primer-pairs per reaction, ranging from 4 to 15-plex (Table 1; Figs. 3 and 4a; Online Resource 1: Table S2). Here, it should be noted that these 290 markers correspond to 270 loci, since the targeted by 20 out of the 25 redesigned SSRs are represented twice in the multiplexes, with the original primers in the first set (Mx01 to Mx34) and with the new primers in the second set (Mx35 to Mx45). In order to test for possible manipulation errors in large-scale studies, all the multiplex PCRs from Mx03 to Mx45 included a “control locus” from Mx01 or Mx02 (Online Resource 1: Table S2). Additionally, in order to verify the consistency of

allele genotyping (“binning” process and allele calling), 20 accessions were duplicated in the validation collection. The control loci included in the multiplexes from Mx03 to Mx45 produced the same genotypes as the initial ones obtained using Mx01 or Mx02. In the same way, identical genotypes were observed for each pair of the duplicated samples at all the loci. New alleles not found neither in the testing panel nor in previous studies were detected for most of the studied loci; even though, allele size overlapping was detected only between the markers VRZAG112 (allelic range 228–260 bp) and VVIN73 (254–267 bp) in Mx01. However, only one rare allele from VRZAG112 (260 bp with 0.029 allele frequency) was implicated, and its electrophoretic profile was clearly distinguishable from the VVIN73 one.

Genetic diversity and quality of the studied SSRs

Out of the 284 markers (the six multilocus ones not included) used to genotype the validation collection with the 45 multiplex PCRs, 247 were successfully amplified in all the 207 individuals. Failure rates lower than 5 % were recorded for 30 markers and larger than 5 % for only 7 markers (Online Resource 1: Table S2). The 264 unique loci (discarding the original versions of the redesigned markers) identified a total of 2,760 alleles, ranging between 1 and 31 per locus (10.45 on average), and most of them (92 %) detected more than four alleles (Fig. 4b; Online Resource 1: Tables S2 and S4). Allelic sizes ranged from 56 to 456 bp (Fig. 4c; Online Resource 1: Table S2). The difference between the smallest and the largest fragment detected at each polymorphic locus ranged between 2 bp (VVIV22, VVIB72 and VVIN78) and 162 bp (VMC5G1-1), with an average of 36.60 bp. Exceptionally,



◀ **Fig. 2** Distribution of the 264 genotyped loci along the grapevine genome. The chromosomal positions (in Mbp) were determined by Blast-search against the 12x.2 *Vitis* genome assembly (Canaguier et al. 2014 <https://urgi.versailles.inra.fr/Species/Vitis/Data-Sequences/Genome-sequences>) using the primer sequences as probes. Vertical bars represent the 19 grapevine chromosomes. Gray segments indicate the chromosomal gaps which were investigated for the development of new markers within located (*in italics*). It should be noted here that markers selection and development of new SSRs to cover the chromosomal gaps were based on the previously available version of the grapevine genome sequence (12x; <http://www.genoscope.cns.fr/>) from which inter-loci distances and ordering have undergone several important changes

the locus Vchr8b amplified fragments ranging from 100 to 455 bp (difference of 355 bp). Online Resource 1: Table S5 shows the allele sizes obtained for all the genotyped SSRs on a subset of 25 well-known and accessible accessions, which could be used as reference samples in other studies. The estimated null allele frequency (F) for the 264 SSRs varied between -0.145 and 0.696 , and 87 % of the loci showed an $F < 0.1$ (Fig. 4d; Online Resource 1: Table S2). The mean values of expected (H_{exp}) and observed (H_{obs}) heterozygosity were 0.699 and 0.658 , ranging from 0.000 to 0.931 and from 0.000 to 0.928 , respectively (Fig. 4e; Online Resource 1: Table S2). The average polymorphic information content (PIC) was 0.665 (0.000 – 0.925), and the average probability of identity for unrelated individuals was 0.15 (0.009 – 1.000) (Fig. 4e; Online Resource 1: Table S2). Beside its genetic diversity, the usefulness of a molecular marker is also determined by the pattern of its amplification profiles, which determines in turn the easiness of allele scoring. The characteristics of the amplification patterns showed by each SSR (stuttering patterns, “+A” effect, low heterozygote peak ratio, presence of artifacts) and a global evaluation of the easiness of scoring of the marker are shown in Online Resource 1: Table S2.

Discussion

SNP markers have been gaining popularity for genetic studies since the last decade due to their high frequency, ability to high throughput automated analysis, and the decrease of their costs (Rafalski 2002; Davey et al. 2011). Even though, micro-satellite markers remain the markers of choice for many genetic studies due to their attractive features (high variability, codominance, transferability, and reproducibility). This is particularly true for studies that require the use of relatively small number of markers or samples. However, the use of large sets of SSR markers is limited due to the high cost and the amount of work required for their analysis, as well as the difficulties of automation. Those limitations could be partially overcome by adopting multiplexing strategies (Butler 2005a, b; Missiaggia and Grattapaglia 2006; Blacket et al. 2012) that would lead to significant savings in time, efforts, and laboratory reagents (Elnifro et al. 2000; Raabová et al. 2010). Empirical results demonstrated that SSR-based multiplexing and multiloading reduced the cost of PCR reagents up to 50 % and the cost of electrophoresis up to 85 % when compared to genotyping based on single-locus PCR (Masi et al. 2003; Merdinoglu et al. 2005). Guichoux et al. (2011) estimated that, even for a moderate number of samples, a 12-plex multiplexing could be eight times cheaper than simplex PCR. Additionally, genotyping errors due to human factor are less likely to occur, as laboratory manipulations are considerably reduced. Two approaches could be used for multiplexing: (1) the amplification of only one or a few markers followed by pooling the products from the individual PCRs prior to electrophoresis (Hall et al. 1996); (2) the joint amplification of all the loci to be loaded in the same electrophoresis run. In this study, we have used the second approach to address the development of an efficient tool that would allow grapevine genome-wide

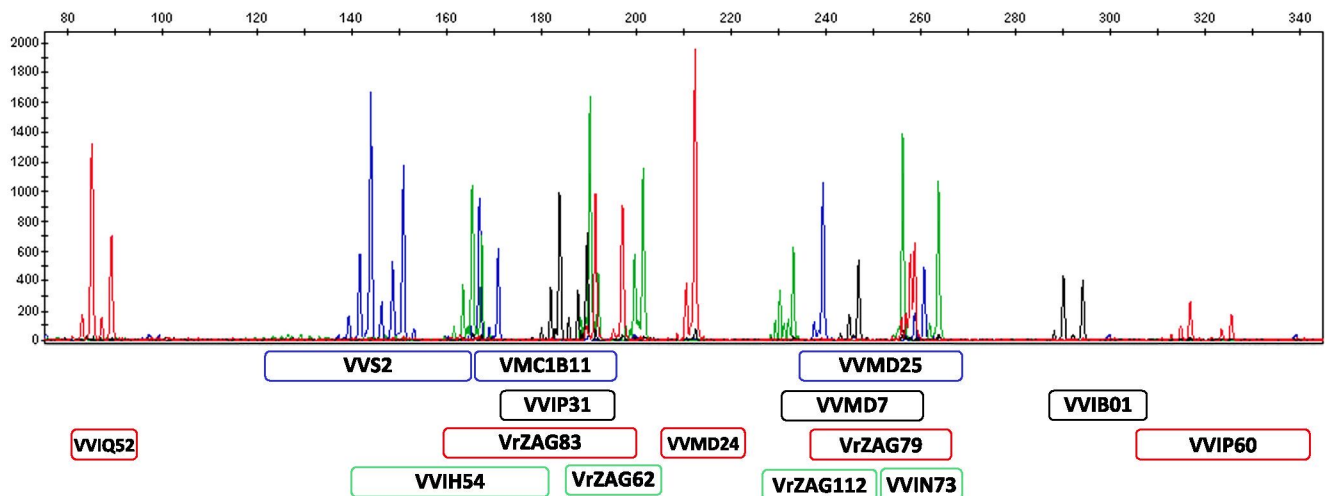
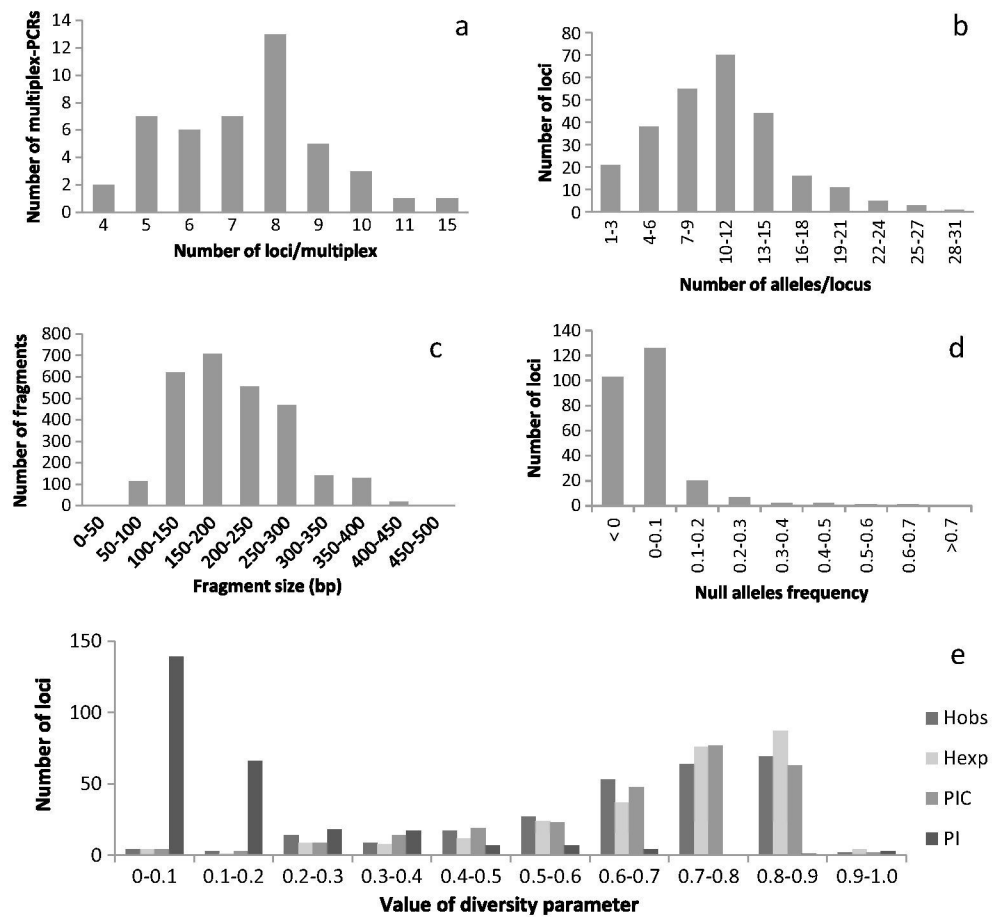


Fig. 3 Example of PCR profile obtained with the multiplex Mx01. FAM-, NED-, PET-, and VIC-fluorescence-labeled SSR markers are indicated in blue, black, red, and green lines, respectively. Length of

the rectangles below the electropherogram represents allelic ranges of each locus

Fig. 4 Distribution of multiplexing level and genetic diversity parameters of the 264 genotyped loci. **a** Number of loci per multiplex PCR. **b** Number of alleles per locus. **c** Size of the 2,760 detected fragments. **d** Estimated frequency of null alleles. **e** Observed (H_{obs}) and expected (H_{exp}) heterozygosity, polymorphic information content (PIC), and probability of identity (PI)



scanning using as fewer as possible PCR runs and electrophoresis tracks. Multiplex PCRs were developed following two main steps: (1) markers selection and organization in multiplex sets; and (2) the optimization of PCR conditions for the proposed multiplexes. In addition, one of the important issues in this work was the establishment of up to seven control points to avoid possible mistakes that could be very difficult to detect in large-scale genotyping projects (see “Data checking” on “Materials and methods”). The approach adopted in this study should be applicable for the development of similar tools in any other species.

Markers selection and multiplex design

SSR markers were selected based on three criteria: genome-wide coverage, high polymorphism and diversity in the fragment size ranges. In spite of the numerous efforts devoted to the development of SSR markers for grapevine genotyping since the early nineties, the number of available SSRs remains limited when compared to other genome sequenced crop species. For instance, whereas 4,109, 8,192, and even 70,732 entries for SSR probes were found at the NCBI for soybean, wheat, and rice, respectively, only 1,079 entries were found for grapevine (<http://www.ncbi.nlm.nih.gov/probe>). In

addition, for most of them (62.5 %) fragment size ranges are concentrated between 150 and 250 bp (Fig. 4c). This limits marker combination possibilities, compromising hence the development of multiplex PCRs with high multiplexing levels (Hill et al. 2009). Moreover, the available information is usually scarce or incomplete, since many of these markers have been developed and/or tested using only a few individuals. For example, in our PCR conditions, five UDV (Di Gaspero et al. 2005), four *Vitis* Microsatellite Consortium (VMC), and one FAM (Huang et al. 2011) SSRs amplified in several accessions more than two fragments in a narrow size interval, resulting in unreliable electrophoretic profiles. This could be related to the presence of multiple binding sites for their respective primers in the grapevine genome. For instance, the design of alternative primers allowed recovering monolocus segregation for four of them. In the same way, three loci described as monomorphic in *V. vinifera* were selected because they were studied in only four (VVIB54; Merdinoglu et al. 2005) or even two accessions of this species (FAM18 and FAM104; Huang et al. 2011), suggesting that more alleles could be detected in a larger sample. Our results revealed that it is unlikely to detect more than one allele at these loci in *V. vinifera*. Nevertheless, those markers, as well as other 21 that detected only two or three alleles (Online

Resource 1: Table S2), could be more informative in genetic backgrounds from other *Vitis* species. For example, the above-mentioned monomorphic markers detected more than one allele when other *Vitis* species were analyzed (Merdinoglu et al. 2005; Huang et al. 2011). On the other side, we found chromosomal regions spanning up to 10 Mbp for which SSR markers had never been described (Fig. 2). Attempting to develop new markers, we identified many microsatellite motifs (up to 153/Mbp) in the investigated regions, but successful primers could be designed for only a few of them. Moreover, PCR failures and/or irregularities in allele inheritance related to the presence of null alleles have been found for nearly half of the 22 new markers. Difficulties with SSR markers development have been already reported in plants (Tero et al. 2006) as well as animal species (Zhang 2004; McInerney et al. 2011), and linked to the presence of repetitive DNA and/or transposable elements. This might be also the case in our study, since most of the SSR gaps were located in low gene density regions of the grapevine genome (Online Resource 2), which had been found substantially complementary to high density of repetitive/transposable elements (Jaillon et al. 2007).

Optimization of multiplex PCRs

The second critical step for the development of the multiplex PCRs consisted in the establishment of suitable thermocycling programs and the adjustment of primers concentrations to obtain a balanced amplification for all the markers included in the same reaction. Touchdown PCR (Don et al. 1991) proved to be highly suitable for multiplex amplification. In fact, only two touchdown-based programs (Tp-A and Tp-B) were enough to run 90 % of the designed multiplex PCRs, which included primers with a wide range of annealing temperatures. As an example, Mx05 includes primers with annealing temperatures ranging from 48 to 67 °C. Touchdown-based thermocycling programs combined with the use of commercial kits optimized for multiplex PCR contribute to save time and labor in the optimization steps when compared to conventional protocols (Masi et al. 2003). In this study, they allowed the successful amplification of up to 15 primer-pairs in the same PCR reaction. As far as we know, this is the highest multiplexing level reached for microsatellite-based grapevine genotyping. Merdinoglu et al. (2005) considered the preferential amplification of small fragments over long fragments for multiplex design, which limited the multiplexing level that could be reached. Our results demonstrated that it is possible to amplify fragments with size differences up to 329 bp in the same reaction; as an example, Mx19 includes 10 loci with fragment sizes ranging from 76 to 401 bp. Nevertheless, an average of 3.8 (between two and eight) optimization assays were needed to achieve balanced amplifications.

Validation of multiplex PCRs

The 45 developed multiplex PCRs were tested and optimized using seven grapevine accessions. In order to verify their validity for genotyping large diversity panels, we used them to genotype a set of 207 accessions representing practically most of the cultivated grapevine genetic diversity. Indeed, this collection was selected from the 1,852 *V. vinifera* accessions maintained at the germplasm bank of El Encín (Alcalá de Henares, Madrid, Spain), using the genotypic dataset of the 26 control loci through the genetic diversity maximization strategy (Le Cunff et al. 2008). Moreover, the main progenitors of grapevine cultivars (Lacombe et al. 2013) are present in the validation collection. The representativeness of this collection was also verified by comparison with two large grapevine germplasm collections: (1) an Italian collection of 745 accessions (Cipriani et al. 2010)—a set of 22 “Vchr” markers identified an average of 9.8 (3 to 21) alleles/locus in that collection, almost the same as in the validation collection (between 3 and 20 with an average of 9.7). (2) A collection of 2,323 *V. vinifera* subsp. *sativa* cultivars conserved at INRA grape repository at Vassal (France) (Laucou et al. 2011)—twenty out of the 26 control loci detected a lower number of alleles (between 5 and 19, with an average of 11.95) in the validation collection than in the Vassal collection, where an average of 16.9 (6 to 36) alleles/locus were identified. However, these differences are mainly related to the presence of uncommon genetic material carrying rare alleles in the larger French collection (2,323 vs 207). For instance, 9 and 12 alleles not detected in our study by the markers VMC4F3 and VVIV67, respectively, had been detected in Laucou et al. (2011), but with frequencies lower than 0.005. These findings point out that, although additional alleles may be detected when studying more distant genetic material, the allelic ranges obtained using the validation collection should represent a good estimation of the actual diversity that can be found in cultivated grapevine. Online Resource 1: Table S5 shows the complete genotypes obtained for 25 reference cultivars at the 264 studied loci that can be used for inter-laboratory comparisons and protocols setting-up.

Null alleles

The presence of null alleles at microsatellite loci can introduce important biases into genetic studies (Callen et al. 1993; Pompanon et al. 2005). Their existence in grapevine have been already demonstrated by sequence analysis (Sefc et al. 1999). Another reliable approach for null allele detection is the analysis of allelic inheritance in family groups (Dakin and Avisé 2004). Using this approach, we identified the probable existence of null alleles in 49 out of the 284 analyzed primer pairs. Most of them (69.39 %) showed estimated null allele frequencies (F) >0.10. However, the absence of null alleles in

the 14 studied trio pedigrees does not discard their presence in the rest of the genotypes. In fact, 22 % of the loci with a moderate to high frequency of null alleles ($0.10 < F < 0.34$) did not show any incompatible genotype among the studied pedigrees. On the other hand, incompatible genotypes were detected for 67.5 % of primer pairs that, after PCR repetition, failed to amplify at least one sample. These results point out that most of the finally recorded PCR failures could be related to the presence of null alleles at homozygote state. Out of the 32 markers showing null alleles for which alternative primers were designed, 18 recovered the normal segregation of alleles, decreasing the estimated null allele frequency in most of the cases (from 0.170 to 0.002 on average), and no amplification failures were noticed for any redesigned primer pair.

Applications of the developed multiplex PCRs

Microsatellite and SNP markers have become in the last years the markers of choice for genetic analyses in grapevine, including genetic mapping and QTL identification (Huang et al. 2012; Doligez et al. 2013; Battilana et al. 2013; Barba et al. 2014), linkage disequilibrium and association analyses (Emanuelli et al. 2010; Barnaud et al. 2010; Cardoso et al. 2012; Vargas et al. 2013), and varietal identification (Myles et al. 2010; Cabezas et al. 2011; Laucou et al. 2011; Migliaro et al. 2012). As in many other woody plant species, genetic mapping in grapevine has been carried out using mainly the double pseudo test-cross strategy (Grattapaglia et al. 1995), which is based on the study of allelic segregation of markers found in heterozygosis in one or both progenitors of an F_1 progeny. The development of new genotyping platforms and technologies has increased the number of SNPs that can be genotyped to hundreds of thousands allowing the construction of highly saturated genetic maps. However, with the currently available information, these markers are not suitable for a direct comparison with the previously published genetic maps and QTLs in grapevine, which are predominantly based on microsatellite markers. The number of markers that can be genotyped using the panel of multiplex PCRs developed in this study is large enough to allow that comparisons. For example, the analysis of the segregation types inferred from the genotypes showed on Online Resource 1: Table S5 points out that in a supposed project involving a mapping progeny derived from a cross between Cabernet Sauvignon and Pinot Blanc would allow to map 235 out of the 264 studied loci, with an average of 12.37 per chromosome (from 7 in chromosome 10 to 17 in chromosome 5). On the other side, linkage disequilibrium in *V. vinifera* expands up to 16 cM when studied using microsatellite markers (Barnaud et al. 2006, 2010). This suggests that a low-density whole-genome genotyping, using the 45 developed multiplex PCRs as genotyping tool, should be useful for QTL detection through association mapping studies.

American and Asian *Vitis* species constitute a valuable source of resistance genes that can be introduced into *V. vinifera* varieties through interspecific breeding programs (Töpfer et al. 2011). The limiting step in this kind of programs is the recovery of the genetic background of the recurrent parent (*vinifera*) by backcrossing. The assistance of the selection process by genome-wide distributed SSRs organized in ready to use multiplex PCRs, such as the developed in this study, has the potential to make considerable savings in time and costs. The efficiency of such a tool could be improved by organizing SSRs from each chromosome in a distinct multiplex PCR as suggested in a previous simulation study (Herzog et al. 2013).

Multiplexes Mx01 and Mx02 allow the study of 26 unlinked markers, including at least one in each of the 19 *V. vinifera* chromosomes. They were designed reorganizing the multiplex PCRs S, A, and B used by Ibáñez et al. (2009) to genotype 376 table grape accessions. These 26 microsatellites have been used to genotype the complete germplasm bank of El Encín (Alcalá de Henares, Madrid, Spain). Twenty of them have been also used through eight multiplex PCRs and three sequencing runs to characterize most of the INRA (France) grape repository (Laucou et al. 2011). Additionally, smaller subsets of the control loci used in this study (including the six OIV and the nine GrapeGen06 SSRs) have been used to characterize grapevine genetic resources from most of the viticultural regions in the world (This et al. 2011). Given the large amount of information available for these markers, Mx01 and Mx02 can be used for a rapid and cost efficient characterization of unexplored grapevine germplasm and short range genetic studies, such as parentage analysis. In this work, we have used Mx01 and Mx02 to certify the identity of each of the extracted DNAs. Moreover, the inclusion of one of the loci amplified with Mx01 and Mx02 as control markers in each multiplex PCR allowed to verify assay reproducibility and sample traceability throughout the study, factors which should be considered by any researcher aware of the consequences of genotyping errors (Pompanon et al. 2005).

Although important cost and time savings might be obtained using the panel of multiplex PCRs presented in this study, a considerable investment in primers labeling will be still required. Economic labeling methods could be used to further decrease the costs. As an example, Blacket et al. (2012) proposed the use of four fluorescently labelled universal primers (one primer for each fluorescent dye) through the three primer PCR approach. This kind of techniques would offer an inexpensive alternative to the commercial synthesizing of custom labeled primers for multiplex genotyping. However, it should be noted that additional optimization steps may be needed because sometimes the three primers approach seem to decrease PCR efficiency (de Arruda et al. 2010).

Conclusions

The genotyping tool developed in this work allows the study of 270 grapevine microsatellite loci using only 45 PCRs and sequencing runs. This represents a significant increase in information gain and time and cost savings when compared to the currently available multiplexes. Moreover, this study allowed to generate a reliable information related to allele size ranges, diversity parameters and presence of null alleles, as well as features of amplification profiles and ease of scoring. This information can be used not only to identify the most suitable loci for future microsatellite-based genetic studies in grapevine (i.e., easy to score, carrying the higher number of alleles, and having the lower estimated frequency of null alleles) but also to design ad hoc multiplexes, combining the most informative loci for specific studies in only a few multiplex PCRs. On the other side, this study has demonstrated that (1) the presence of null alleles is common in grapevine microsatellite markers and, hence, it should be considered for any population genetics or linkage disequilibrium-based studies; (2) large chromosomal regions for which SSR markers have never been developed still exist in the grapevine genome, and it seems that this is related to a high density of repetitive/transposable elements in these regions.

Acknowledgments This study was financially supported by project RTA2009-00149-00-00 from Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA, Spain) and project FP08-AL-03 (Comunidad de Madrid, Sapin). BZ was supported by a PhD scholarship from INIA. MA was supported by a PhD scholarship from Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario (IMIDRA). JAC worked under Doctor's recruiting and hiring contract of INIA-CCAA.

Data archiving statement Primer sequences of the redesigned and newly developed SSR markers are provided in Online Resource 1: Table S2, and they are being submitted to the NCBI. Genotypic data for the 284 SSRs in 25 grapevine accessions are provided in Online Resource 1: Table S5.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Arroyo-García R, Martínez-Zapater J (2004) Development and characterization of new microsatellite markers for grape. *Vitis* 43:175–178
- Ban Y, Mitani N, Hayashi T et al (2014) Exploring quantitative trait loci for anthocyanin content in interspecific hybrid grape (*Vitis labruscana* × *Vitis vinifera*). *Euphytica* 198:101–114. doi:10.1007/s10681-014-1087-3
- Barba P, Cadle-Davidson L, Harriman J et al (2014) Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. *Theor Appl Genet* 127:73–84. doi:10.1007/s00122-013-2202-x
- Barnaud A, Lacombe T, Doligez A (2006) Linkage disequilibrium in cultivated grapevine, *Vitis vinifera* L. *Theor Appl Genet* 112:708–716. doi:10.1007/s00122-005-0174-1
- Barnaud A, Laucou V, This P et al (2010) Linkage disequilibrium in wild French grapevine, *Vitis vinifera* L. subsp. *silvestris*. *Heredity* 104:431–437. doi:10.1038/hdy.2009.143
- Battilana J, Lorenzi S, Moreira FM et al (2013) Linkage mapping and molecular diversity at the flower sex locus in wild and cultivated grapevine reveal a prominent SSR haplotype in hermaphrodite plants. *Mol Biotechnol* 54:1031–1037. doi:10.1007/s12033-013-9657-5
- Blacket MJ, Robin C, Good RT et al (2012) Universal primers for fluorescent labelling of PCR fragments—an efficient and cost-effective approach to genotyping by fluorescence. *Mol Ecol Resour* 12:456–463. doi:10.1111/j.1755-0998.2011.03104.x
- Bouquet A (2011) Grapevines and viticulture. In: Adam-Blondon AF, Martínez-Zapater JM, Kole C (eds) *Genetics, genomics and breeding of grapes*. Science Publishers, New Hampshire, pp 1–29
- Bowers JE, Dangl GS, Vignani R, Meredith CP (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). *Genome* 39:628–633
- Bowers JE, Dangl GS, Meredith CP (1999) Development and characterization of additional microsatellite DNA markers for grape. *Am J Enol Vitic* 50:243–246
- Bryda EC, Riley LK (2008) Multiplex microsatellite marker panels for genetic monitoring of common rat strains. *J Am Assoc Lab Anim Sci* 47:37–41
- Butler JM (2005a) Forensic DNA typing protocols. *Methods Mol Biol* 297:53–65. doi:10.1385/1592598676
- Butler JM (2005b) Constructing STR multiplex assays. *Methods Mol Biol* 297:53–66
- Cabezas JA, Ibáñez J, Lijavetzky D et al (2011) A 48 SNP set for grapevine cultivar identification. *BMC Plant Biol* 11:153. doi:10.1186/1471-2229-11-153
- Callen DF, Thompson AD, Shen Y et al (1993) Incidence and origin of “null” alleles in the (AC)_n microsatellite markers. *Am J Hum Genet* 52:922–927
- Canaguier A, Scalabrín S, Le Paslier M-C, et al. (2014) Improvement of the grapevine genome assembly. *Plant and Animal Genome Conference (PAGXXII)*
- Cardoso S, Lau W, Eiras Dias J et al (2012) A candidate-gene association study for berry colour and anthocyanin content in *Vitis vinifera* L. *PLoS One* 7:e46021. doi:10.1371/journal.pone.0046021
- Cipriani G, Marrazzo MT, Di Gaspero G et al (2008) A set of microsatellite markers with long core repeat optimized for grape (*Vitis* spp.) genotyping. *BMC Plant Biol* 8:127. doi:10.1186/1471-2229-8-127
- Cipriani G, Spadotto A, Jurman I et al (2010) The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor Appl Genet* 121:1569–1585. doi:10.1007/s00122-010-1411-9
- Cipriani G, Di Gaspero G, Canaguier A et al (2011) Molecular linkage maps: Strategies, resources, and achievements. In: Adam-Blondon A-F, Martínez-Zapater J-M, Kole C (eds) *Genetics, genomics and breeding of grapes*. Science Publishers, New Hampshire, pp 68–89
- Correa J, Mamani M, Muñoz-Espinoza C et al (2014) Heritability and identification of QTLs and underlying candidate genes associated with the architecture of the grapevine cluster (*Vitis vinifera* L.). *Theor Appl Genet* 127:1143–1162. doi:10.1007/s00122-014-2286-y
- Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* 93:504–509. doi:10.1038/sj.hdy.6800545

- Davey JW, Hohenlohe PA, Etter PD et al (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499–510. doi:10.1038/nrg3012
- De Arruda MP, Gonçalves EC, Schneider MPC et al (2010) An alternative genotyping method using dye-labeled universal primer to reduce unspecific amplifications. *Mol Biol Rep* 37:2031–2036. doi:10.1007/s11033-009-9655-7
- Decroocq V, Favé MG, Hagen L et al (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. *Theor Appl Genet* 106:912–922. doi:10.1007/s00122-002-1158-z
- Di Gasparo G, Cipriani G, Marrazzo MT et al (2005) Isolation of (AC)_n-microsatellites in *Vitis vinifera* L. and analysis of genetic background in grapevines under marker assisted selection. *Mol Breed* 15:11–20. doi:10.1007/s11032-004-1362-4
- Doligez A, Bertrand Y, Farnos M et al (2013) New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). *BMC Plant Biol* 13:217. doi:10.1186/1471-2229-13-217
- Don RH, Cox PT, Wainwright BJ et al (1991) “Touchdown” PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* 19:4008
- Drašnarová A, Krak K, Vít P et al (2014) Cross-amplification and multiplexing of SSR markers for *Ambrosia glutinosa* and *A. incana*. *Tree Genet Genomes* 10:865–873. doi:10.1007/s11295-014-0727-z
- Duchêne E, Butterlin G, Dumas V, Merdinoglu D (2012) Towards the adaptation of grapevine varieties to climate change: QTLs and candidate genes for developmental stages. *Theor Appl Genet* 124:623–635. doi:10.1007/s00122-011-1734-1
- Elnifro EM, Ashshi AM, Cooper RJ, Klapper PE (2000) Multiplex PCR: optimization and application in diagnostic virology. *Clin Microbiol Rev* 13:559–570. doi:10.1128/CMR.13.4.559-570.2000
- Emanuelli F, Battilana J, Costantini L et al (2010) A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). *BMC Plant Biol* 10:241. doi:10.1186/1471-2229-10-241
- Gouesnard B (2001) MSTRAT: an algorithm for building germ plasm core collections by maximizing allelic or phenotypic richness. *J Hered* 92:93–94. doi:10.1093/jhered/92.1.93
- Grattapaglia D, Bertolucci FL, Sederoff RR (1995) Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross strategy and RAPD markers. *Theor Appl Genet* 90:933–947. doi:10.1007/BF00222906
- Grzeskowiak L, Costantini L, Lorenzi S, Grando MS (2013) Candidate loci for phenology and fruitfulness contributing to the phenotypic variability observed in grapevine. *Theor Appl Genet* 126:2763–2776. doi:10.1007/s00122-013-2170-1
- Guichoux E, Lagache L, Wagner S et al (2011) Current trends in microsatellite genotyping. *Mol Ecol Resour* 11:591–611. doi:10.1111/j.1755-0998.2011.03014.x
- Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. *Heredity* 101:5–18. doi:10.1038/hdy.2008.35
- Hall JM, LeDuc CA, Watson AR, Roter AH (1996) An approach to high-throughput genotyping. *Genome Res* 6:781–790
- Herzog E, Töpfer R, Hausmann L et al (2013) Selection strategies for marker-assisted background selection with chromosome-wise SSR multiplexes in pseudo-backcross programs for grapevine breeding. *Vitis* 52:193–196
- Hill CR, Butler JM, Vallone PM (2009) A 26plex autosomal STR assay to aid human identity testing. *J Forensic Sci* 54:1008–1015. doi:10.1111/j.1556-4029.2009.01110.x
- Huang H, Lu J, Ren Z et al (2011) Mining and validating grape (*Vitis* L.) ESTs to develop EST-SSR markers for genotyping and mapping. *Mol Breed* 28:241–254. doi:10.1007/s11032-010-9477-2
- Huang Y-F, Doligez A, Fournier-Level A et al (2012) Dissecting genetic architecture of grape proanthocyanidin composition through quantitative trait locus mapping. *BMC Plant Biol* 12:30. doi:10.1186/1471-2229-12-30
- Ibáñez J, Vargas AMA, Palancar M et al (2009) Genetic relationships among table-grape varieties. *Am J Enol Vitic* 60:35–42
- Jaillon O, Aury J-M, Noel B et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449:463–467. doi:10.1038/nature06148
- Jewell MC, Frere CH, Prentis PJ et al (2010) Characterization and multiplexing of EST-SSR primers in *Cynodon* (*Poaceae*) species1. *Am J Bot* 97:99–101. doi:10.3732/ajb.1000254
- Kalia RK, Rai MK, Kalia S et al (2010) Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177:309–334. doi:10.1007/s10681-010-0286-9
- Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 16:1099–1106. doi:10.1111/j.1365-294X.2007.03089.x
- Karaagac E, Vargas AM, Andrés MT et al (2012) Marker assisted selection for seedlessness in table grape breeding. *Tree Genet Genomes* 8:1003–1015. doi:10.1007/s11295-012-0480-0
- Lacombe T, Boursiquot JJ-M, Laucou V et al (2013) Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera* L.). *Theor Appl Genet* 126:401–414. doi:10.1007/s00122-012-1988-2
- Lauco V, Lacombe T, Dechesne F et al (2011) High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theor Appl Genet* 122:1233–1245. doi:10.1007/s00122-010-1527-y
- Le Cunff L, Fournier-Level A, Laucou V et al (2008) Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. *sativa*. *BMC Plant Biol* 12:31. doi:10.1186/1471-2229-8-31
- Lefort F, Kyvelos CJ, Zervou M et al (2002) Characterization of new microsatellite loci from *Vitis vinifera* and their conservation in some *Vitis* species and hybrids. *Mol Ecol Notes* 2:20–21. doi:10.1046/j.1471-8286.2002.00130.x
- Liu L, Wu Y (2012) Development of a genome-wide multiple duplex-SSR protocol and its applications for the identification of selfed progeny in switchgrass. *BMC Genomics* 13:522. doi:10.1186/1471-2164-13-522
- Lodhi MA, Reisch BI (1995) Nuclear DNA content of species, cultivars, and other genera of the *Vitaceae*. *Theor Appl Genet* 90:11–16. doi:10.1007/BF00220990
- Martins WS, Lucas DCS, de Neves KFS, Bertioli DJ (2009) WebSat—a web software for microsatellite marker development. *Bioinformatics* 3:282–283
- Masi P, Zeuli P, Donini P (2003) Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.). *Mol Breed* 303–313.
- McInerney CE, Allcock AL, Johnson MP et al (2011) Comparative genomic analysis reveals species-dependent complexities that explain difficulties with microsatellite marker development in molluscs. *Heredity* 106:78–87. doi:10.1038/hdy.2010.36
- Mejía N, Soto B, Guerrero M et al (2011) Molecular, genetic and transcriptional evidence for a role of VvAGL11 in stenospermocarpic seedlessness in grapevine. *BMC Plant Biol* 11:57. doi:10.1186/1471-2229-11-57
- Merdinoglu D, Butterlin G, Bevilacqua L et al (2005) Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. *Mol Breed* 15:349–366. doi:10.1007/s11032-004-7651-0
- Migliaro D, Morreale G, Gardiman M et al (2012) Direct multiplex PCR for grapevine genotyping and varietal identification. *Plant Genet Resour* 11:182–185. doi:10.1017/S1479262112000433
- Missiaggia A, Grattapaglia D (2006) Plant microsatellite genotyping with 4-color fluorescent detection using multiple-tailed primers. *Genet Mol Res* 5:72–78

- Moreno-Sanz P, Loureiro MD, Suárez B (2011) Microsatellite characterization of grapevine (*Vitis vinifera* L.) genetic diversity in Asturias (Northern Spain). *Sci Hortic* 129:433–440. doi:10.1016/j.scienta.2011.04.013
- Myles S, Chia J-M, Hurwitz B et al (2010) Rapid genomic characterization of the genus *Vitis*. *PLoS One* 5:e8219. doi:10.1371/journal.pone.0008219
- Myles S, Boyko ARA, Owens CCL et al (2011) Genetic structure and domestication history of the grape. *Proc Natl Acad Sci U S A* 108:3530–3535. doi:10.1073/pnas.1009363108
- Patil DT, Bronner MP, Portier BP et al (2012) A five-marker panel in a multiplex PCR accurately detects microsatellite instability-high colorectal tumors without control DNA. *Diagnostic Mol Pathol* 21:127–133. doi:10.1097/PDM.0b013e3182461cc3
- Pelsy F (2010) Molecular and cellular mechanisms of diversity within grapevine varieties. *Heredity* 104:331–340. doi:10.1038/hdy.2009.161
- Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* 6:847–859. doi:10.1038/nrg1707
- Postolache D, Leonarduzzi C, Piotti A et al (2013) Transcriptome versus genomic microsatellite markers: highly informative multiplexes for genotyping *Abies alba* Mill. and congeneric species. *Plant Mol Biol Report* 32:750–760. doi:10.1007/s11105-013-0688-7
- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1:215–222. doi:10.1016/1360-1385(96)86898-1
- Raabová J, Hans G, Risterucci A-M et al (2010) Development and multiplexing of microsatellite markers in the polyploid perennial herb, *Menyanthes trifoliata* (*Menyanthaceae*). *Am J Bot* 97:e31–e33. doi:10.3732/ajb.1000084
- Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol* 5:94–100. doi:10.1016/S1369-5266(02)00240-6
- Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends Genet* 9:275–280. doi:10.1016/0168-9525(93)90013-8
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365–386
- Sayama T, Hwang T-Y, Komatsu K et al (2011) Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean. *DNA Res* 18:107–115. doi:10.1093/dnares/dsr003
- Schlötterer C (2004) The evolution of molecular markers—just a matter of fashion? *Nat Rev Genet* 5:63–69. doi:10.1038/nrg1249
- Scott KD, Eggler P, Seaton G et al (2000) Analysis of SSRs derived from grape ESTs. *Theor Appl Genet* 100:723–726. doi:10.1007/s001220051344
- Sefc KM, Regner F, Turetschek E et al (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome* 42:367–373
- Sefc KM, Pejić I, Maletić E et al (2009) Microsatellite markers for grapevine: tools for cultivar identification & pedigree reconstruction. In: Roubelakis-Angelakis KA (ed) *Grapevine molecular physiology & biotechnology*. Springer, Dordrecht, pp 565–596
- Smith JR, Carpten JD, Brownstein MJ et al (1995) Approach to genotyping errors caused by nontemplated nucleotide addition by Taq DNA polymerase. *Genome Res* 5:312–317. doi:10.1101/gr.5.3.312
- Tero N, Neumeier H, Gudavalli R, Schlötterer C (2006) *Silene tatarica* microsatellites are frequently located in repetitive DNA. *J Evol Biol* 19:1612–1619. doi:10.1111/j.1420-9101.2006.01118.x
- This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes. *Trends Genet* 22:511–519. doi:10.1016/j.tig.2006.07.008
- This P, Zapater JM, Péros J-P, Lacombe T (2011) Natural variation in *Vitis*. In: Adam-Blondon AF, Martínez-Zapater JM, Kole C (eds) *Genetics, genomics and breeding of grapes*. Science Publishers, New Hampshire, pp 30–67
- Thomas MR, Scott NS (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theor Appl Genet* 86:985–990. doi:10.1007/BF00211051
- Thomas MR, Matsumoto S, Cain P, Scott NS (1993) Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. *Theor Appl Genet* 86:173–180. doi:10.1007/BF00222076
- Töpfer R, Ludger H, Eibach R (2011) Molecular breeding. In: Adam-Blondon AF, Martínez-Zapater JM, Kole C (eds) *Genetics, genomics and breeding of grapes*. Science Publishers, New Hampshire, pp 160–185
- Torregrosa L, Fernandez L, Bouquet A et al (2011) Origins and consequences of somatic variation in grapevine. In: Adam-Blondon AF, Martínez-Zapater JM, Kole C (eds) *Genetics, genomics and breeding of grapes*. Science Publishers, New Hampshire, pp 68–92
- Vargas AM, Fajardo C, Borrego J et al (2013) Polymorphisms in *VvPel* associate with variation in berry texture and bunch size in the grapevine. *Aust J Grape Wine Res* 19:193–207. doi:10.1111/ajgw.12029
- Velasco R, Zharkikh A, Troggio M et al (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS One* 2:e1326. doi:10.1371/journal.pone.0001326
- Wagner AP, Creel S, Kalinowski ST (2006) Estimating relatedness and relationships using microsatellite loci with null alleles. *Heredity* 97:336–345. doi:10.1038/sj.hdy.6800865
- Welter L, Ms G, Zyprian E (2011) Basics of grapevine genetic analysis. In: Adam-Blondon AF, Martínez-Zapater JM, Kole C (eds) *Genetics, genomics and breeding of grapes*. Science Publishers, New Hampshire, pp 137–159
- Ye J, Coulouris G, Zaretskaya I et al (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13:134. doi:10.1186/1471-2105-13-134
- Zhang D-X (2004) Lepidopteran microsatellite DNA: redundant but promising. *Trends Ecol Evol* 19:507–509. doi:10.1016/j.tree.2004.07.020